‘Khoisan’ sibling terminologies in historical perspective: a combined anthropological, linguistic and phylogenetic comparative approach[1]

Gertrud Boden, Tom Güldemann & Fiona Jordan

1. Introduction

In this paper we combine regional anthropological comparison, historical linguistics and phylogenetic comparative methodology (PCM) in addressing the historical relationships between the languages of the three South African ‘Khoisan’[2] families, Kx’a, Tuu and Khoe-Kwadi (see Güldemann, introduction, this volume). Since the data on extinct Kwadi are insufficient, this language had to be excluded, so that we will hereafter only refer to Khoe. Generally, the demonstrable linguistic relationships within Kx’a (Heine & Honken 2010), Tuu (Güldemann 2005), and Khoe (Vossen 1997) imply original family-specific sibling terminologies with relevant lexemes as part of the proto-languages used within a social culture of the proto-societies (cf. Murdock 1949: 346f; Elmendorf 1961: 365; Jordan 2011: 299). By looking first at the ways how siblings are grouped into kin classes and secondly how sibling terms as lexical items are linguistically related we develop a number of hypotheses for proto-terminologies, contact scenarios and trajectories of change to be submitted to PCM probability tests. By trying to detect signals of genealogical or contact relationships we hope to contribute to the reconstruction of pre-historical processes in the Kalahari Basin, including testing hypotheses found in the previous literature, among them the claim about a deep structural unity of Khoisan kinship systems.

From the early days of social anthropology, kinship terminologies have been discussed as a source of information on ancient population dynamics (Morgan 1871). Murdock (1968) considered sibling classifications to be particularly suitable for confirming or refuting hypotheses of language relationships (cf. also Dziebel 2007). While sibling terminologies have indeed been shown to be useful for reconstructing culture history in different parts of the world (Firth 1970; Epling, Kirk & Boyd 1973; M. Marshall 1984; Hedican 1986; Blust 1994; Jordan 2011), they were not in the focus of previous comparison in the Khoisan area (Barnard 1988, 1992; Ehret 2008).

There are, however, advantages when analyzing sibling terms historically. Compared to terms for other basic kin, they are less likely to be similar due to independent development steered by universal factors (cf. nursery forms for “mother” and “father”). Compared to terms for more distantly related kin like cousins, their classification is less likely to be affected by changing norms of marriage and residence. Finally, they have been claimed to be less affected by borrowing or encroachment from terms for more distant relations (Ehret 2008). Other reasons to focus here on sibling terminologies are more pragmatic: First, sufficient documentation of sibling terms is also available for at least some of the now extinct Khoisan languages. Second, the documented sibling terminologies show a sufficient degree of structural variation in order to apply PCM (see §4).

Many anthropologists took a purely structural approach to the reconstruction of sibling classifications by looking only at social dimensions such as relative age, sex of referent, and

[1] This research, as part of the European Science Foundation EUROCORES Programme EuroBABEL, was supported by the British Economic and Social Research Council and the Deutsche Forschungsgemeinschaft, for which we would like to express our thanks. We also acknowledge support from the Max-Planck-Gesellschaft.

[2] We do not assume ‘Khoisan’ to be a language family when using the term for convenience.

relative sex (cf. Kroeber 1909; Murdock 1968).[3] Recent phylogenetic work also relied solely on such information (Jordan 2011: 308). Disregarding the history of sibling terms themselves has been criticized by linguists (see the comments by Blust, Clark and Chowning to M. Marshall 1984). The limited geographical scope and number of languages in our study allowed us to take an innovative approach in developing a tailor-made typology for the Kalahari Basin area, recording not only the social dimensions but also the linguistic means (lexical stems, modifiers, derivational devices, and grammatical elements) by which they are expressed.
The sibling terminologies analyzed here are taken from a variety of sources such as dictionaries, ethnographic descriptions, and data from fieldwork made available through personal communication. In general we gave preference to ethnographically informed descriptions, because understanding of the classificatory meanings is expected to be deeper; wordlists and dictionaries were consulted for linguistic accuracy. Different orthographies were subjected to a general unified transliteration which is explained in §3.1. Due to insufficient documentation not all historically attested languages are included in our analysis. Those investigated are listed in Table 1; this should be compared with the full inventory given in Güldemann (introduction, this volume). In spite of a number of gaps, especially in the East Kalahari Khoe group, our data include more speech communities than earlier comparisons (e.g., Barnard 1988, 1992) and has a fuller coverage of different primary groups.

2. Systems of sibling classification

We first give a survey of Khoisan sibling terminologies from an anthropological perspective. The overview is found in Table 1 where a basic type of sibling classification is assigned to each language-specific system, using the same letters as in Murdock (1968).[4] The six basic types relevant for the discussion are as follows; all but the first are attested in the sample languages today:

Type A: generic sibling term only = no other social dimension
Type B: relative age only
Type C: relative age + sex of referent for elder sibling
Type D: relative age + sex of referent
Type E: sex of referent only
Type F: relative sex - unlabeled subtype: + relative age for same sex sibling

INSERT Table 1: A survey of sibling classifications across Khoisan[5]

From a typological perspective, the Khoisan sibling classifications are all 'mainstream'. None contains more than four terms, as do 85% of Murdock's (1968) world sample of 800 societies. They are also consistent with Nerlove & Romney's (1967) established universal principles, viz.: a) if a sex-of-referent-distinction is only made among one pair of siblings it will be among elder siblings, and b) if a relative-age distinction is only made among one pair of siblings it will be the same sex siblings.

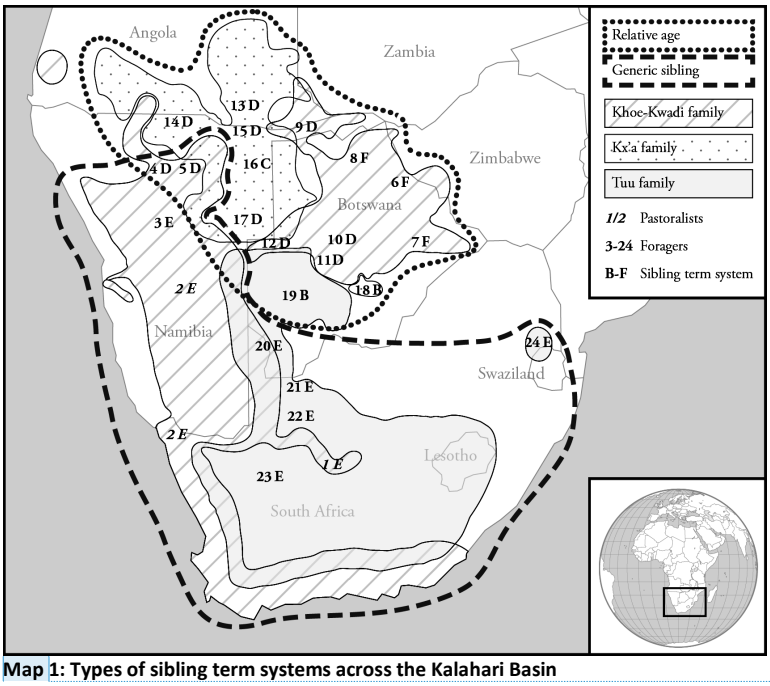
---

[3] Another possible dimension is sex of speaker which, however, is not relevant in any of the known Khoisan languages, apart from the fact that observing a relative-sex distinction (= same sex vs. opposite sex sibling) indirectly reflects sex of speaker.

[4] We also add information on a group's subsistence, because this and correlating features of social organization may influence a particular sibling term type (see discussion below).

[5] The abbreviations for sibling terms used in Table 1 and tables or figures further below are as follows: B brother, e elder, G sibling, ms man speaking, os opposite sex, ss same sex, ws woman speaking, y younger, Z sister.

A second general observation is that the number of five basic types in a relatively small set of 24 partly related language varieties is high and suggests a considerable degree of historical dynamics, either due to a high rate of internal change or to cultural borrowing over a sufficiently long common history in the area.



Map 1: Types of sibling term systems across the Kalahari Basin

Commented [FJ1]: If at all possible, I would recommend that:
- 1E, 24E, 12D and 18E sit "outside" their circles with a line drawn to indicate where they belong.
- If that's not possible, the illustrator should nudge the font closer together
- Lesotho sit within the boundaries or at least have a grey box behind the text so that the border does not cut a letter in half
- Botswana & Namibia to have a white box behind so the grey hatching doesn't intersect with text

Finally, the distribution of most types correlates little with genealogical language groups, which holds on the level of primary families as well as family-internal sub-groups. Instead, they show geographical patterns, often across genealogical boundaries, as seen in Map 1. Sibling classifications of Type E are found in 8 languages belonging to the !Ui and Lower Nossob groups of Tuu and the Khoekhoe branch of Khoe; the area (indicated by a dashed line) straddles the southern half of the Kalahari Basin area and, due to the expansion of the Khoekhoe group, portions in the northwest. A second quite compact area is found in the northern part of the Kalahari Basin, formed by 10 languages of Type D from the Khoe and Kx'a families. Type B is found in 2 unrelated languages in a small central area: Taa of the Tuu family and ǂHoan of the Kx'a family. Type F is attested in 3 languages in the north-eastern part of the Kalahari Basin, viz. Shua, Ganadi and ǀXokhwe - all from Kalahari Khoe, although not in the same sub-branch. North Juǀ'hoan around Tsumkwe and Dobe has uniquely Type C.[6] Given that Types B, C, D, and F all encode relative age to some extent a yet larger area with this feature

[6] Ironically, the anthropological documentation of San has been strongly biased toward this particular group (cf. Barnard 1992; Takada 2008), so that their culture became regarded as being representative of the whole Ju group if not San societies in general.

(indicated by a dotted line) can be identified in the northern Kalahari Basin. The areal and genealogical distributions are discussed in more detail in §5.

3. Lexical reconstruction of sibling terminology

3.1 Preliminaries

In a second step, we attempt a first reconstruction of family-specific sibling terms by means of historical-comparative linguistic methodology, in which proto-forms with a certain sound shape and meaning are hypothesized to have been present in an earlier language stage. Since language-specific terms are not transcribed by scholars in a homogeneous way, we transliterated them into a unified orthography which makes them comparable. This orthography applies the following rules based on regularities across the languages at issue. Lexical stems largely comply with just three phonotactic patterns: $C(C)_1V_1C_2V_2$, $C(C)_1V_1V_2$, and $C(C)_1V_1N$. These involve striking parallels between consonants and their coarticulations on the one hand and phonation types of vowels on the other hand. Moreover, the conventions are oriented towards practical orthographies which try to avoid special symbols and diacritics (cf. Dickens 1994, Güldemann 1998a, Naumann forthcoming). Notations which are not straightforward from general linguistic practice are as follows: c = (alveo)palatal voiceless fricative; g = voicing before click and uvular C_1; n = alveolar nasal [n] in C_2 but nasalization before click C_1 and after V_2; nn = word-final alveolar nasal; h = aspiration after C_1 and breathiness after V_1; hh = glottal fricative in C_1 cluster; j = (alveo)palatal voiced fricative; q = uvular stop in C_1 cluster and pharyngealization after V_1; x = posterior fricative in C_1 cluster; ’ = glottalization after C_1 and V_1 and before nasal (gesture in) C_1; ’’ = glottal stop in C_1 cluster. Lexical stems are generally bimoraic so that a single V in an original transcription is normally changed into VV. Moreover, when necessary, vowels are given in their assumed underlying phonological form, notably *i(i)* as /ai/. Tone marking has been removed because of incomplete information.

A general caveat for the following discussion is that it can only be a first attempt at reconstructing parts of the kinship terminology in the three families. The hypotheses to be presented here need to be corroborated by careful family-specific research.

3.2 The Kx’a family

We start our analysis with the Kx’a family. Table 2 provides the transliterated sibling terminology of five Ju dialects and the isolated ǂHoan language.

Table 2: Transliterated sibling terminology of Kx’a

Meaning	**ǂHoan**	**Ju**				
		Southeast		**North-central**		**North**
		South Juǀ’hoan	North Juǀ’hoan	Lloyd’s !Xuun[7]	Ekoka !Xuun	Angola !Xuun
eZ	*ki si(i)*	*!ui*	*!ui*	*ǁkhui*	*ǁui*	*ǁui*
yZ	*ǁam*	*tsisi(n)*	*sin, tshin*	*tsasing*	*tahng*	*tacing*
eB	*ki si(i)*	*!o*	*!o*	*ǁkò*	*ǁo*	*ǁo*
yB	*ǁam*	*tsi(-ma)*	*sin, tshin*	*sing*	*cing*	*tcing*

The linguistic comparison within Kx’a is primarily concerned with Ju, because ǂHoan forms are unique. For “elder sibling”, Proto-Ju forms can be reconstructed without major problems: *ǃǃui “elder sister” and *ǃǃoo “elder brother” (cf. already Sands 2010). Both series involve a well

[7] “Lloyd’s !Xuun” is used here for a data set which was recorded by L. Lloyd and whose geographical origin cannot yet be determined conclusively. According to Lionnet (p.c.), it most likely belongs in the North-central group within Sands’ (2010) dialect classification.

documented sound change from a retroflex to an alveolar click in south-eastern dialects and to a lateral click in all other dialects which don't retain the original retroflex *ǃǃ (ibid: 89-91). The different click accompaniment in Lloyd's *ǁkhui* "elder sister" must be viewed as an unexplained idiosyncrasy of the source. In view of the general root pattern CV(C)V the "elder brother" form is likely to have been *ǃǃoo with a sequence of like vowels, although this feature is not transcribed in any source used. Since both reconstructions start with the same click and a similar vowel, viz. /u/ and /o/, it is conceivable that there was actually a single root *ǃǃV^{back} for "elder sibling", from which one of the sex-specific forms was derived by a vocalic suffix that was incorporated by the root, viz. *ǃǃoo-i to *ǃǃui or *ǃǃui-o to *ǃǃoo.

The forms in Ju for "younger sibling" pose more intricate problems. The analysis is facilitated by also evaluating them together because they display a number of similarities as well as systematic differences. It is crucial to observe first that /ts/, /tc/, /s/, /c/, and /h/ are voiceless, phonetically stronger consonants which normally occur in Khoisan languages only in root~morpheme-initial position. Hence, more than one such consonant in a lexeme indicates at least diachronically more than one root~morpheme. Except for North Ju|'hoan, which lacks a distinction between different sex forms, the number of these consonants regularly distinguishes "younger sister" from "younger brother": while the latter has only one the former has two. Moreover, the single root in the simpler "younger brother" forms is identical or at least similar with the second root in the complex "younger sister" forms. Such a more abstract root can be analyzed as conveying the meaning "younger sibling", which is modified by a prefix for rendering the derived female form. The abstract root for "younger sibling" can be reconstructed plausibly as *tsing. In its simplex form for "younger brother" it becomes *tsin with a nasalized vowel in Southeast Ju which in turn gives rise to *sin* in North Ju|'hoan via weakening of /ts/ to /s/, and to *tsi(-ma)* in South Ju|'hoan via loss of nasalization (the suffix -*ma* marks diminutives and is optional). In the other Ju forms, the initial consonant of *tsing is weakened to *sing*, palatalized to *tcing*, or weakened and palatalized to *cing*. The final segment *-tsing in the more complex "younger sister" forms behaves similarly, except that the degree of weakening of the initial /ts/ is taken one step further: one finds *-h(i)ng* < *-cing/-sin(g)* < *-tcing/*-tsing; this is not unexpected for a word-medial consonant. The initial segments in the "younger sister" forms can be derived from an original prefix comprising an alveolar consonant followed by the vowel /a/. The /i/ in *tsi-* of South Ju|'hoan is the result of assimilation to the following root. The initial consonant is reconstructed here as /th/ because this is the best candidate to yield both /ts/ via weakening and /t/ via loss of aspiration. In North Ju|'hoan "younger sibling" seems to be expressed by two forms *sin* and *tshin*. In view of the situation in other dialects it is conceivable that the sex distinction between reconstructed *tsing and *t(h)a-tsing became neutralized but both forms were retained whereby *tshin* would be the reflex of *t(h)a-tsing via an intermediate form *tsi-sin, as still found in South Ju|'hoan.

Table 3: Reconstructed sibling terminology and system types in Kx'a

Meaning	**ǂHoan**	**Ju**
Elder sibling = eG	*ki si(i)*	*ǃǃV^{back}
Younger sibling = yG	*ǁam*	*tsing
Sex (with elder sibling)	-	*-V (*ǃǃoo-I > *ǃǃui or *ǃǃui-O > *ǃǃoo)
Female (with younger sibling)	-	*t(h)a-
Type	B	*B > *D

Table 3 summarizes all (reconstructed) linguistic elements involved in sibling terms of the Kx'a family. No Proto-Kx'a lexemes can be established. However, there may at least have been systemic-structural similarities between ǂHoan and Ju if the earliest stage in Ju was indeed of Type B with *ǃǃV^{back} "elder sibling" vs. *tsing "younger sibling", and the sex-of-referent distinction is secondary. However, these reconstructed sex-indicating derivations, while

suggestive for formal and system-internal reasons, are not without problems and must not be taken already as robust reconstructions. First, both processes are so far not attested with other words let alone are transparent with respect to their possible origin. That is, they would be old and idiosyncratic with these kinship terms (see a parallel case in Tuu to be mentioned in §3.3). Second, their different form, and even more so, their different morphotactic status, initial modification for female with "younger sibling" vs. final sex modification with "elder sibling", requires an explanation, which is so far unavailable. One possible solution is that the two distinct processes emerged in different historical layers of the language.

3.3 The Tuu family

We proceed in the following with lexical reconstructions in the second family, Tuu. Table 4 presents the transliterated forms of the languages surveyed.

Table 4: Transliterated sibling terminology of Tuu

Meaning	Taa	Lower Nossob		!Ui		
		ǀHaasi	ǁ'Auni	N‖ng	ǀXam	‖Xegwi
eZ	*Ɵxaa*	*(n) ‖xae*	*‖xae*	*‖aa-xae*	*‖aa-xae*	*‖aa-ke*
yZ	*ǂqx'ann*	*(n) ‖xae*	*‖xae*	*‖aa-xae*	*‖aa-xae Ɵwaa*	*‖aa-ke*
eB	*Ɵxaa*	*(n) ‖aa-si*	*‖aa-si*	*‖aun*	*‖aan*	*‖aa*
yB	*ǂqx'ann*	*(n) ‖aa-si*	*‖aa-si*	*‖aun*	*‖aan Ɵwaa*	*‖aa*

The three basic sub-groups of this family are neatly reflected in Table 4 by (partly) distinct terminology. The language complex Taa is internally homogeneous and one can identify in the extensive dialect data available to us two linguistic forms, namely *Ɵxaa* "elder sibling" and *ǂqx'ann* "younger sibling" without any other distinction.

Since !Ui is internally more diverse and presumably historically older, the forms for sibling terminology display more variation. Linguistic reconstruction is nevertheless relatively straightforward. Once all forms are transliterated in a unified way one can identify a basic root *‖aa "sibling". This root together with the female suffix *-xae renders "sister". The male counterpart "brother" is, except for ‖Xegwi, brought about by changing the final segment of the neutral stem through a different vowel quality and/or nasalization. It is probable that this process is also due to an earlier suffix which was incorporated by the root, because both morphological processes are also found with a few other kinship terms in !Ui (cf. Güldemann 2005: 22). A distinction regarding relative age cannot be reconstructed for the proto-language. In some modern languages it is conveyed by optionally attaching a diminutive derivation, as in ǀXam, or by relative-type clauses conveying "big/great" and "small", as in ‖Xegwi.

Finally, the Lower Nossob forms require some more discussion, because the linguistic material generally has deficiencies which make it more difficult to interpret. The pattern is that the original forms in the available sources are reminiscent of those in !Ui but differ from them in their initial segments. We argue that these differences are only superficial, due to a partly inappropriate analysis and transcription as well as phonological processes that are also discernible in related !Ui languages, notably ǀXam. The relevant sibling terms in Table 4 are already adapted accordingly.

The ǀHaasi form for "brother" in the original source is *‖ŋa: si*; it differs from Proto-!Ui *‖aa- by an additional nasal gesture on the click. It needs to be taken into account, though, that kinship terms are inherently related to a 'possessor' and when elicited are mostly delivered by speakers with a possessor pronoun, usually of first person singular. This pronoun in Tuu has the form of a syllabic nasal /n/, /ŋ/, /m/ depending on its following consonant. In fast speech and with limited familiarity of the language, a researcher may well perceive such a segment before a simple click merely as click nasalization and write it accordingly as a phonetic gesture after the click symbol. We thus assume that *‖ŋa: si* is actually *n ‖aasi* "my brother", parallel to

n ǁxai, likewise meaning "my sister". This is supported by Bleek's (1937) l'Auni forms for "brother" which are simply *ǁka:si* or *ǁkas* (= *ǁaas(i)*).
The original |Haasi form *ǁxai* "sister" in turn should not be compared directly with *ǁaa-xae* in |Xam and N‖ng but the variant |Xam forms given in (1). Here, the shorter forms on the left apparently arise from a fast pronunciation of the longer forms on the right, two of which are recorded as the principle variants given in Table 4.

(1) "sister" *ǁxai* (= *ǁxae*) < *ǁaa-xae*
"younger sister" *ǁxa ʘpwa* (= *ǁxaa ʘʷaa*) < *ǁaa-xae ʘʷaa*
"younger sisters" *ǁxᶜaukẹ̊n ʘpu̥onni* (= *ǁxau-knn ʘʷani*) < *‖aa-xu-knn ʘʷani
(Bleek 1956: 631)

Bleek (1937) also gives the variation recorded in (1) for |'Auni, viz. *ǁka:xe* (=*ǁaa-xae*) vs. *ǁxei* (=*ǁxae*), although here it cannot be excluded that the long form is a borrowing from neighboring N‖ng, due to the intimate contact between the two languages. In any case, we can assume on the basis of the |Xam data alone that |Haasi *ǁxai* and l'Auni *ǁxei* also derive from an underlying *‖aa-xae. In summary one can thus identify for Proto-Lower Nossob the same root *‖aa "sibling", which was specified for sex by different suffixes: *-xae for female (singular) and *-si for male (singular).
A summary of the possible reconstructions in the three Tuu subgroups is given in Table 5. As in Kx'a, it is not possible to arrive at a unified Proto-Tuu system either in structural or lexical terms, in particular because Taa stands out with words that are entirely different in both form and meaning. Proto-!Ui and Proto-Lower Nossob, however, are quite comparable in having a Type-E system with a single sibling term that is transparently modified by sex-specific suffixes. Since these markers are obviously secondary, one can conclude that the original system in terms of Murdock (1968) was of Type A, which does not mark any social dimension.

Table 5: Reconstructed sibling terminology and system types in Tuu

Meaning	**Taa**	**Lower Nossob**	**!Ui**
Elder sibling = eG	*ʘxaa*	-	-
Younger sibling = yG	*ǂqx'ann*	-	-
Sibling = G	-	*‖aa	*‖aa
Female	-	*-xae	*-xae
Male	-	*-si	*-Ṽ
Type	B	*A > *E	*A > *E

3.4 The Khoe family

Finally, we try to reconstruct sibling terminology in the Khoe family. Since the data surveyed are the most extensive and complex, we will proceed with linguistic reconstruction in the two major branches separately and combine the results at the end. Table 6 starts with the transliterated data on languages of the Khoekhoe group.

Table 6: Transliterated sibling terminology of Khoekhoe

Meaning	**South**	**North**			
	!Ora	Nama	Damara	Hai‖om	ǂAakhoe
eZ	*!aan-s*	*!aan-(sa)-s*, rarely *kai-s*	*kuin-sa-s*	*ausi-s*	*ai-s*
yZ	*!aan-s*	*!aan-(sa)-s*	*kuin-sa-s*	*!aan-s*	*ano-s*
eB	*!aan-b*	*!aan-(sa)-b*	*kuin-sa-b*	*abudi-b*	*ai-b*
yB	*!aan-b*	*!aan-(sa)-b*	*kuin-sa-b*	*!aan-b*	*ano-b*

The systems code sex of referent and partly relative age. All languages mark the first feature by means of person-gender-number (PGN) suffixes, *-b* for masculine and *-s* for feminine. While this marking is obligatory, it is a grammatical feature holding for all nouns of these languages

and is thus not dedicated to the terminology system as such. The expression of relative age is not lexicalized in !Ora, Nama, and the Damara variety presented here, which are accordingly of Type E (in the first two, "elder" can be conveyed by using the adjective *kai* "big", including bare *kai-s* in Nama; it goes back to an identical Proto-Khoe form, cf. Voßen 1997: 445). ǂAakhoe, Hai‖om, and some Damara dialects not dealt with here display the relative-age distinction in root lexemes and are thus of Type D, which unites all Namibian non-pastoralist Khoekhoe groups. A second important area of diversity has a similar distribution: the three non-pastoralist groups largely have lexical roots distinct from *!aan* in pastoral Nama and !Ora.
The existence of Khoekhoe varieties in Namibia is commonly associated with the expansion of pastoralist groups from South Africa (Budack 1986, Vedder 1997). Since the non-pastoralist Damara, Hai‖om, and ǂAakhoe would then have resulted from "Khoekhoeization" processes, they are unlikely to be closer to the proto-system. So Proto-Khoekhoe is reconstructed here with a single root *!aan "sibling" with grammatical marking of sex and optional indication of relative age.

Table 7: Transliterated sibling terminology of Kalahari Khoe[8]

Meaning	**West**					**East**	
	Naro	G‖ana	G|ui	Khwe	|Xokhwe	Shua	Ganadi
eZ (ws)	*kai-sa*	*kya-xo-si*	*tya-xo-si*	*ta-ci-hɛ*	*da-si*	*tya-hu*	*tca-xo*
yZ (ws)	*!uin-sa*	*daba-xo-si*	*dyiba-xo-si*	*dama-ci-hɛ*	*dama-ce*	*dama-hon*	*dama-xo*
eB (ws)	*kai-ba*	*kya-xo-m*	*tya-xo-bi*	*ta-ci-ma*	*guin-ke*	*k'uin-k'e*	*kuin-ke*
yB (ws)	*!uin-ba*	*daba-xo-m*	*dyiba-xo-bi*	*dama-ci-ma*	*guin-ke*	*k'uin-k'e*	*kuin-ke*
eZ (ms)	*kai-sa*	*kya-xo-si*	*tya-xo-si*	*ta-ci-hɛ*	*guin-ke*	*k'uin-k'e*	*kuin-ke*
yZ (ms)	*!uin-sa*	*daba-xo-si*	*dyiba-xo-si*	*dama-ci-hɛ*	*guin-ke*	*k'uin-k'e*	*kuin-ke*
eB (ms)	*kai-ba*	*kya-xo-m*	*tya-xo-bi*	*ta-ci-ma*	*da-si*	*tya-hu*	*tca-xo*
yB (ms)	*!uin-ba*	*daba-xo-m*	*dyiba-xo-bi*	*dama-ci-ma*	*dama-ce*	*dama-hon*	*dama-xo*

Table 7 contains the comparative sibling terminology in Kalahari Khoe. Like in Khoekhoe, sex of referent is often marked by gender suffixes (*-sa*, *-si*, *-hɛ* for feminine and *-ba*, *-bi*, *-m(a)* for masculine). Although only present in the West group except |Xokhwe, these PGN suffixes can be reconstructed further back because they go back to Proto-Khoe *-sV and *-bV (cf. Voßen 1997, Güldemann 2004). Their absence on elicited sibling terms in eastern languages is in line with an overall cline: grammatical PGN-marking is rare in the (north)east of the family and only becomes more regular or even obligatory further (south)west (cf. Güldemann 2004, forthcoming).
There are two other linguistic elements which are found across the entire Kalahari Khoe group, except Naro. These are *t(y)a~kya "elder" and *daba~dama "younger" (the deviant *dyiba* in G|ui is the result of palatalization from *daba via *dyaba). The relevance of both reconstructable forms differs, however, in the modern systems: they only apply to same sex siblings in East Kalahari Khoe and |Xokhwe but are general in the other relevant West Kalahari Khoe languages.
The greatest diversity in the group exists in the lexical sibling roots themselves. In order to identify possible proto-forms, isolated lexemes should be excluded first. On the one hand, this concerns Naro *kai* "elder sibling" which, like in Khoekhoe, seems to derive from the Proto-Khoe adjective *kai "big" but has turned into a nominal base. On the other hand, there are *ci* "sibling" in Khwe and the presumably related *si/ce* in |Xokhwe; their origin is uncertain (see §5.1, footnote 18, for a possible hypothesis).

[8] |Xokhwe is a data set recorded by Heinz (n.d., field notes) in ???Khwai and grouped with Buga which is closely related to Khwe. Ganadi is a speech variety recorded by Westphal in the ???Tuli block in the east of Botswana and thus belongs to East Kalahari Khoe.

Naro *!uin* "younger sibling" also seems to be unique in Kalahari Khoe. It is more likely, however, that the form is related to other sibling terms in Khoe. These are *guin-ke*, *k'uin-k'e*, and *kuin-ke* "opposite sex sibling" in Kalahari Khoe, and *kuin* or *!uin* "sibling" in Damara. Our very preliminary reconstruction is *!uin-(k(')e) whose exact meaning cannot yet be determined; the candidates are "sibling" or "opposite sex sibling".[9]
A similar semantic problem holds for the robust reconstruction *xo (changing in Shua to *hu/hon*). In East Kalahari Khoe it means "same sex sibling", while in G‖ana and G|ui of West Kalahari Khoe it is a generic root "sibling". Recall from §2 that the opposite vs. same sex distinction in parts of Kalahari Khoe, expressed via the simplex forms *kuin-ke and *xo, is unique among the Khoisan languages surveyed here.
All possible reconstructions for Khoe are summarized in Table 8. Only the PGN-marking provides some Proto-Khoe forms encoding the sex-of-referent dimension; the feature is, however, grammatically rather than structurally induced. There are no other obvious systemic or linguistic commonalities across the two branches.

Table 8: Reconstructed sibling terminology and system types in Khoe

Meaning	**Khoekhoe**	**Kalahari**
(Same-sex) sibling = (ss)G	-	*xo
(Opposite-sex) sibling = (os)G	-	*!uin-(k(')e)
Sibling = G	*!aan	-
Elder = e	*kai 'big'	*t(y)a~kya
Younger = y	-	*daba~dama
Feminine	*-sV	(*-sV)
Masculine	*-bV	(*-bV)
Type	*E	*F or *D

Summarizing the above first attempt at lexical reconstructions of sibling terms in the three language families, the major finding reiterates that of the anthropological survey in §2: the lexical stock encountered in this domain is considerable, not just across the major families but even within them. This fact requires historical explanation, not only with respect to the validity of the doubtful Khoisan unit but also the dynamics within the three more securely established language families.

4. Phylogenetic comparative analyses

In order to answer precisely this sort of research questions, viz. how linguistic and cultural features change over time, and what forms those features took in the past, scholars in the field of cultural evolution have used in recent years phylogenetic comparative methodology (PCM) from evolutionary biology. By combining knowledge of the relationships between ethnolinguistic groups in the form of family trees (phylogenies), and tracing the evolution of cultural features on those trees with probabilistic statistical models, we can perform a sort of virtual archaeology on aspects of culture and social life that leave no material trace. So we apply in a third step such an approach to our kinship data from the Kalahari Basin populations. We demonstrate how it can both quantify our findings, and in particular arbitrate the central question of the relationship between the three families.

[9] Current research is underway to clarify this intricate set of terms on the basis of more data from Kalahari Khoe. Regular sound changes at least indicate an initial click consonant *!. According to Voßen (1997: 329, 331), this would have changed to *k in East Kalahari Khoe. For Damara, Table 6 displays *kuin*, a recent transcription by Haacke and Eiseb (2002). Vedder's (1923) earlier work on Damara actually has *!uin* which conforms to the expected sound correspondences in this variety. The status of the final element *-k(')e in Kalahari Khoe is still unclear. See §5.4 for some more data and discussion.

We first provide a brief sketch of the principles behind PCM.[10] These methods are statistical techniques that allow the researcher to understand how the current diversity in some feature domain, such as kin terms, evolved through time. They use phylogenetic trees as a scaffold of history on which to track the evolution of features through time. Note that this is not tree-building or -inference itself: these methods USE trees, they do not create them. Given the controversy of using a branching phylogeny to represent the history of ethnolinguistic groups, discussed by Borgerhoff Mulder (2006), it needs to be stressed that there are simple pragmatic reasons for taking this approach. No other statistical method can visualise and properly control for the effects of shared ancestry on feature diversity. Known to anthropologists as Galton's Problem, the fact that languages or populations share features simply because they are historically related is problematic and means that we cannot tally instances of a feature and assume that these counts will be independent for statistical purposes.

PCM avoids this problem by identifying the *independent evolutionary events* themselves and has been developed to address a wide range of evolutionary questions. Many of these questions are central for the historical social sciences, and PCM provides the best quantitative approach currently on offer to issues like the following. Does change in one feature drive change in another? Which features do languages share because of contact and which are due to shared inheritance? Can we infer the nature of features in the past? What model of change best describes a feature?

In Figure 1 we show how a phylogenetic approach incorporates information about historical relationships when determining feature diversity. Panel A shows nine languages with one of two features (black and white) plotted 'geographically'. In a majority sense (6:3), and using age-area inference where features at the edge of a distribution are considered older, black would be considered ancestral. In Panel B historical information is taken into account and features are mapped onto a phylogeny of the languages. This phylogeny reveals that there were only two independent changes to black (grey boxes), rather than six instances. Accordingly, white is more likely to have been ancestral for the group. In general, a PCM will use (a) trees, (b) the feature data and (c) different models of change between states (e.g. black to white and vice versa), to infer both the ancestral state at each node and the best model of change.

[10] For a detailed non-technical description with applications to kinship see Jordan (in press). Nunn (2011) is an introduction to PCM for anthropologists and linguists. Relevant literature reviews are Gray et al. (2007), Mace & Jordan (2011), and Levinson & Gray (2012).

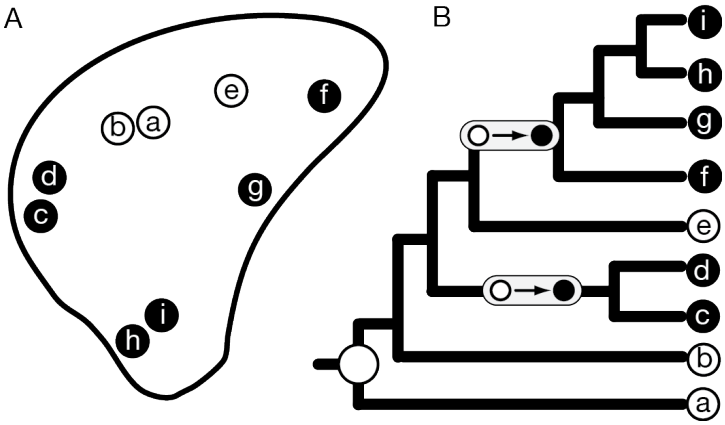


Figure 1: Phylogenetic approaches to understanding feature diversity

We now apply PCM in order to tackle such questions with respect to Khoisan sibling terminology. This requires the use of a phylogeny upon which to model the kinship data, even though the notion that Kx'a, Tuu, and Khoe-Kwadi constitute a single family descending from some putative Proto-Khoisan is nowadays a marginal position. Basic vocabulary has been used in phylogenetic tree-inference for a number of language families (e.g. Austronesian: Gray et al. (2009); Indo-European (Gray & Atkinson 2003); Bantu (Holden 2002), Semitic (Kitchen et al 2008), but such cognate-coded data was not available to us. To overcome this difficulty, we developed a novel approach that took into account both the integrity of the three separate families and what is known and not known about both the internal relationships of the languages and their relative 'ages'. It also allowed us to model the evolution of sibling terms across these three families in a single framework. This meant that the same sorts of cultural processes that act on kinship systems could be assumed to apply across this area, but the languages were not constrained to form a unitary family.

We converted the current language classification into a tree using Mesquite 2.72 (Maddison & Maddison 2011), standard software for the creation and manipulation of phylogenies that is routinely used by evolutionary biologists. Languages were nested together according to the classificatory subdivisions to create a tree that was not fully bifurcating, i.e. it contained 'multifurcations'. We call this the 'base tree'. Because PCM requires fully bifurcating trees for calculating a model of evolution, we developed a procedure that created a large sample of fully bifurcating trees that, taken together, incorporate all of the ambiguity represented by the base tree. This way we are able to test our hypotheses about sibling term evolution over all the thousands of possibilities in history rather than relying on a single classification. This procedure subgrouped languages randomly wherever there was ambiguity.

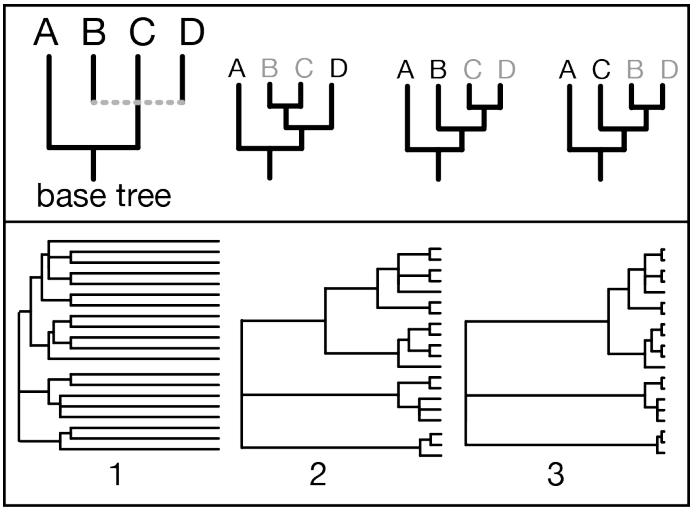


Figure 2: Resolving an ambiguous tree (top) and varying the branch lengths (bottom)

The top panel in Figure 2 shows a sample base tree on the left in which languages B, C, & D are ambiguously related (the dotted line). This could be because they constitute a dialect chain, because they are closely related varieties, or simply because branching relationship cannot be determined. The three trees to the right show all the ways in which this three-way branching can be resolved.

Our base tree, shown in Figure 3 and incorporating 20 of the 24 languages[11] in Table 1, was imported into the R 2.15.1 programming environment (R Core Team 2012) in order to use the phylogenetics packages ape 3.05 (Paradis, Claude & Strimmer 2004) and geiger 1.31 (Harmon et al. 2007). Here, we made 1000 copies[12] of the base tree and applied the "multi2di" function from the geiger package to each tree. This function randomly resolves the multifurcations in each tree in the following way. Where relations between any three languages or subgroups are flat/ unresolved, this function randomly chooses two branches to cluster together to the exclusion of the third. In this way the 1000-tree sample will (as a whole) still contain all the uncertainty about classification, but each tree will be a fully-nested bifurcating tree.

We also cover three scenarios differing according to the amount of change along different lineages, by creating phylogenies with different branch lengths that represent the amount of change (not time) that has taken place. The bottom inset of Figure 2 visualises such variation under three evolutionary scenarios. In Scenario 1, most of the evolutionary changes take place along the terminal (language-specific) branches, implying that recent change is predominant, and that languages share very recent ancestry. In Scenario 2, the branch lengths that lead back from the tips of the tree are roughly equivalent in length to the number of splitting events (nodes), with branches becoming longer as we go further back. Here, changes will be fairly evenly distributed along branches. In Scenario 3, the terminal branches are very short, while the basal branches are comparatively longer. This implies that most evolutionary change takes

[11] We have excluded 4 languages whose addition would not have contributed to the analysis, because they are "sisters" of other languages with the exact same type and cognate lexemes.

[12] We determined by means of standard phylogenetic calculations that this number is sufficient for the data at hand.

place early on and that the three branches are more or less independent. If Kx'a, Tuu, and Khoe-Kwadi are indeed separate families, then the kin-term data will, we predict, fit Scenario 2 or even better Scenario 3, rather than Scenario 1.

We conducted two sets of analyses because the sibling terminologies can be modelled in terms of (1) their system types (cf. Table 1) and (2) the linguistic devices used to express the relevant dimensions. We used a maximum likelihood method called MultiState implemented in the PCM software BayesTraits (Pagel, Meade & Barker 2004) to infer the best model of sibling type evolution. In this procedure, the ancestral state at each node, and the rates of change between each type, are estimated simultaneously over many thousands of combinations and then over each of the trees. Even with this small data set, these calculations are prohibitive for a human analyst who would not be able to evaluate all these possibilities at once, so that computational approaches are necessary (see Pagel, Meade & Barker 2004; Jordan 2011, in press) for more technical details on the method). One advantage of this method is that changes along branches from one state to another (from Type B to D, for example) are allowed to happen at different rates. Thus, if a feature is borrowed, rather than inherited, our results will indicate a higher rate of change in that particular state of the feature.

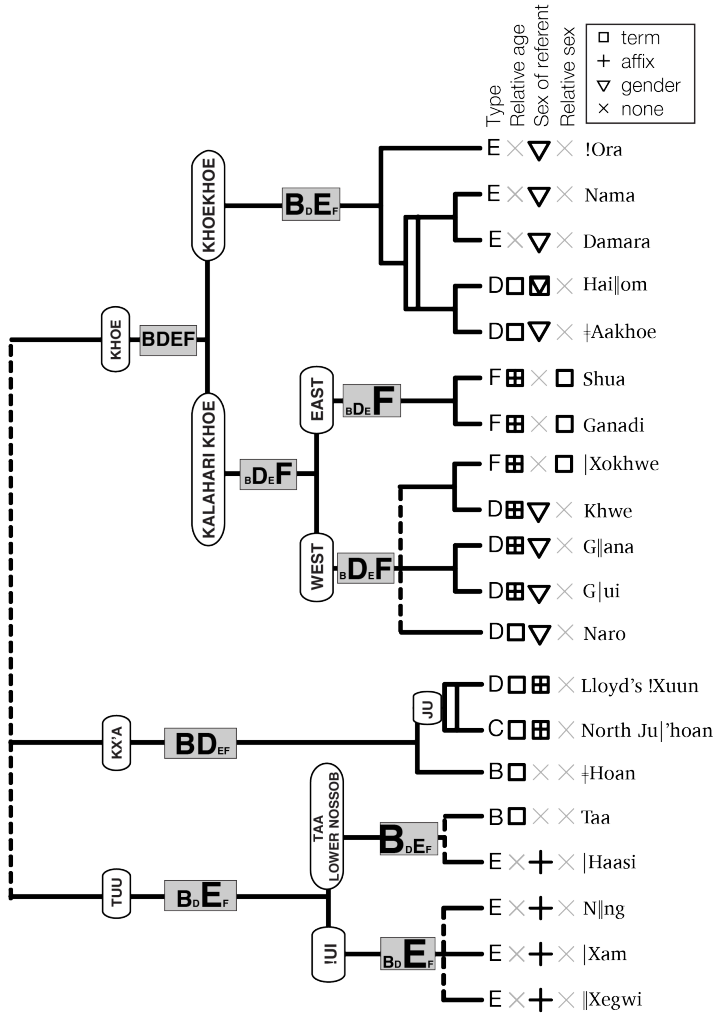


Figure 3: Base tree of the Kalahari Basin languages for PCM analysis

Figure 3 shows the base tree of the 20 languages analyzed. The tree is annotated with subgroups where dotted lines indicate unknown relationships (if any); double lines indicate dialect chains or closely related language varieties. Languages are labelled according to four features: (1) Sibling classification type (cf. Table 1), (2) relative age, (3) sex of referent, (4) relative sex. The symbols indicate the linguistic expression(s) of a given dimension (some languages have more than one) and X indicates that the dimension is not expressed. Internal branches of the tree are annotated with inferred ancestral states of sibling classification types,

Commented [FJ2]: Weird, it's there now. Footnote 13

Commented [T3]: Fiona, you made the following comment: I did use both term and affix in those cases. It doesn't have any impact on the Type reconstructions. Looking at the results it has made no difference to the Kalahari Khoe results (this is still is reconstructed as relative age with affix). For Ju it has made it a more conservative test for sex of referent but because the analysis considers all the possibilities it would still have come out with the same uncertainty about whether sex of referent was present in the Kx'a group. I have added a footnote to that effect.
I did not find any new footnote in your file?

where the size of the letter indicates the robustness of the inference (see Table 9 for precise figures).

We inferred the best model of change, and the ancestral states at each of the nodes on the tree[13]. As shown in §2, the surveyed Khoisan languages display five different classificatory types (B–F). A type that is only expressed in a single language reduces the statistical power of the PCM with only 20 languages, so we recoded Type C in North Ju|'hoan as Type D on the basis of their formal similarities in the linguistic analysis.[14] MultiState gives a score of best-fit (the likelihood) for how each of the three scenarios is compatible with the data. Recall that Scenario 1 implied recent shared ancestry between the families, Scenario 3 virtually-independent lineages, and Scenario 2 a middle ground. Scenario 3 had the highest median likelihood over the 1000 trees (Lh -17.47), followed by 2 (Lh -17.63), then 1 (Lh -20.9). Using a likelihood-ratio test, we can state that Scenario 1 was a significantly worse fit to the data than Scenarios 2 and 3 (p=0.011 and 0.009 respectively, df = 1). In other words, the data do not support a scenario where the three lineages share recent ancestry, in line with current thinking about language relationships in the Khoisan domain.

Figure 4 shows a graphic version of the maximum likelihood model for Scenario 3 (the best scenario). Arrow width is equivalent to the median rate of change between two sibling system types (median values adjacent to arrows). That is, the thicker the arrow, the more often that change is likely to have happened in history. Only seven of a possible 12 transition types are necessary here; half of the possible transition changes had a median rate of zero, reflected by the absence of an arrow.

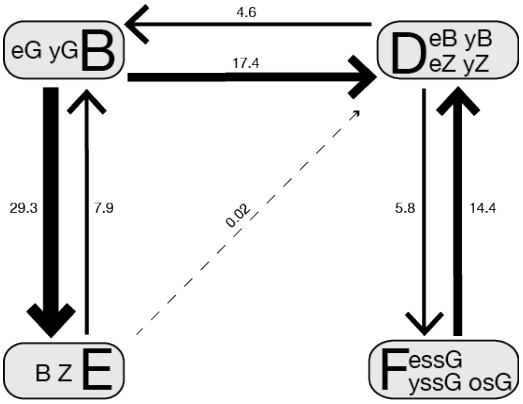


Figure 4: The best model of sibling term type evolution

[13] In the phylogenetic analysis, both term and affix were coded as expressing the dimensions of relative age (in Kalahari Khoe) and sex of referent (in Ju) in comparison to the detailed linguistic analysis. This had no effect on the estimation of Type, and made no difference for the reconstruction of Kalahari Khoe (relative age with affix). For Ju, it is hard to estimate what linguistic device is used to express sex of referent (see Table 9).

[14] With only 20 languages and five sibling types, the Multistate method is unable to reach stable solutions because there are too many possibilities to consider. Coding North Ju|'hoan alternatively as "missing" would have allowed it to take all possible states; these solutions were equivalent to coding it as Type D.

Three of the seven changes make good theoretical and logical sense in line with the literature on changes in kinship terminology suggesting that gains and losses of dimensions tend to happen in a stepwise fashion (cf. §2 and 5.1). These are the changes between Type B and D, requiring only the gain or loss of sex of referent, and the change between E and D that involves the gain or loss of a relative age distinction. Between D and F two dimensions are involved (sex of referent and relative sex). Although we do not investigate the sex-of-speaker dimension, it could be invoked here. For example, "elder brother" in Type D as spoken by a man would be equivalent to "elder same-sex sibling". To go from D to F, then, would simply involve losing the age distinction for a man's sister or a woman's brother. According to the "stepwise" assumption prevalent in kinship theory, a change between B and E would be seen as more problematic, as this involves a wholesale shift (i.e. more than a single change in one dimension) between a relative-age system (B) and a sex-of-referent system (E) (predominantly towards E). Our most-likely model suggests that this change is the most frequent one in the Kalahari Basin, and we discuss some cultural and historical reasons why this might be a salient change in §5.2.

While the fit of the data was not as good, similar models of sibling type evolution were obtained for the two other evolutionary scenarios, so the model of sibling classification does not seem to be absolutely dependent on the type of branch-length scenario. The one point of difference was that for Scenarios 1 and 2, the model contained marginally higher rates of change between Types E and D. We further examined the transitions between different sibling classifications by looking at the rates of gain and loss of the three dimensions. In separate analyses, we examined the gain and loss of the dimensions separately (as opposed to the types). Both relative age and sex of referent were gained at three times the rate they were lost, although sex of referent was more prone to change overall, as demonstrated by the relative rates (relative age 6.3 to 2.3, sex of referent 40.9 to 13.2). Relative sex showed the opposite pattern, as rates of loss were five times higher than gain (15.6 to 2.7).

Finally, the estimates of the ancestral types and expressed social dimensions in various ancestor nodes of Khoisan are given in Table 9. The nodes are followed by the language varieties considered in a group (see Figure 3). The probabilities range from 0 (certain absence) to 1 (certain presence). In the 2nd column for types we indicate in bold when there is robust evidence for a type's presence at that node, viz. where a probability is >.6. In the 3rd column for social dimensions, probabilities are followed with the probable linguistic devices that were inferred to express the dimension.

Table 9: Ancestral state estimates for types and social dimensions

NODE: languages	**Probability of type**	**Probability of social dimension**[15]
KX'A: ǂHoan, Ekoka !Xuun, North Ju|'hoan	B = .48, D = .45 E = .02, F = .03	Relative age = .97 (term) Relative sex = .12 Sex of referent = .45
TAA-LOWER NOSSOB: |Haasi, Taa	**B = .70**, D = 0 E = .30, F = 0	Relative age = .27 (term) Relative sex = .03 Sex of referent = .31 (affix)
!UI: ‖Xegwi, |Xam, N‖ng	B = .13, D = 0 **E = .86**, F = 0	Relative age = .001 Relative sex = .07 Sex of referent = .66 (affix)
TUU: TAA-LOWER NOSSOB + !UI	B = .48, D = .45 E = .02, F = .03	Relative age = .01 Relative sex = .34 Sex of referent = .5 (affix)
KHOEKHOE: !Ora, Damara, Nama, ǂAakhoe, Hai‖om	B = .33, D = .02 **E = .64**, F = .01	Relative age = .13 Relative sex = .34

[15] Inferring the presence or absence of the dimension was done with binary-coded data. A separate multistate analysis inferred the linguistic nature of the dimension (affix, term, grammatical gender, none), and we report the feature(s) that had the highest probability.

		Sex of referent = .51 (gender suffix)
WEST KALAHARI KHOE: Naro, G|ui, G‖ana, Khwe, |Xokhwe	B = .05, D = .47 E = .01, F = .47	Relative age = .91 (affix/term) Relative sex = .49 Sex of referent = .5 (gender suffix)
KALAHARI KHOE: WEST KALAHARI KHOE + Shua, Ganadi	B = .09, D = .37 E = .01, F = .53	Relative age = .75 (affix) Relative sex = .57 (term) Sex of referent = .5 (gender suffix)
KHOE KHOEKHOE + KALAHARI KHOE	B = .25, D = .27 E = .22, F = .26	Relative age = .52 (affix/term) Relative sex = .5 Sex of referent = .5 (gender suffix/term)

The results in Table 9 highlight a difference between phylogenetic and linguistic comparative methodology. Notably, PCM can infer a proto-feature even if no descendant shows it, because it uses a model of change that applies to all the data, i.e. here the entire set of languages, even calculating in potential past diversity obliterated in the course of time. To take an example, in the Khoekhoe group, relative sex is inferred with .34 probability for the proto-stage, even though no modern language has this feature. In other words, PCM uses the data and models of probable feature changes overall, and because it is probabilistic, it allows proto-stages to take states that are not found in the present.

Commented [T4]: Fiona, since the .52 for relative age of the last version had to be removed. So I took this example. Can this stay, or is there a possible error?

Commented [FJ5]: Absolutely fine, and makes more sense now it is a lower probability too!

5. Historical trajectories in Khoisan sibling terminologies

5.1 Preliminaries

The assumption within comparative anthropological approaches (cf. Kuper 1987: 8f; Hage 1999) is that individual types of kinship classifications can be conceived of as regular transformations of each other, also within a culture area, and the ultimate aim is to order them chronologically and derive a plausible model which leads back to a historically shared proto-structure.[16] Our data and the PCM analysis are in line with the current assumption among specialists that Khoisan should so far not be treated as a genealogical entity. Accordingly, we do not seek a historical scenario deriving all attested systems from a single ancestral one. However, the emergence of the diversity within the three secure families, Kx'a, Tuu and Khoe should and can be evaluated. The fact that not a single proto-system of sibling terms could be identified so far implies that after the proto-stages individual subgroups and languages have been subject to considerable change.

Recall another crucial point from §2: basic system types show a bias in geographical rather than genealogical distribution, notably according to just two large areas in the Kalahari Basin (cf. Map 1): one in the north with a variety of types but the common denominator of encoding relative age, and another in the south (and northwest) with Type E whose additional hallmark is the existence of a generic sibling term. Namibian North Khoekhoe varieties of the pastoral Nama and some closely related Damara groups aside, the two areas are complementary to each other. Such a pattern indicates that historical dynamics were partly steered by language contact across genealogical boundaries.

Contact-induced change can concern both the systemic organization of lexemes defining the basic type and their linguistic substance in the form of loan words. Both possibilities seem to be relevant in the Kalahari Basin. This can be illustrated best with the linguistic items conveying the relative-age distinction in the northern area.

Table 10 presents first all primary lexemes with this meaning, which reflect its being an integral part of the terminology systems.

Table 10: (Reconstructed) forms for "elder/younger sibling" across families

Meaning	ǂHoan (Kx'a)	Taa (Tuu)	Naro (Khoe)	Ju (Kx'a)	ǂAakhoe (Khoe)	Hai‖om (Khoe)

[16] The phylogenetic methodology in §4 takes a similar approach but has some differences.

Elder sibling	*ki si(i)*	*Ǝxaa*	*kai-*	*!|V^{back}	*ai-*	*ausi-s, abudi-b*
Younger sibling	*‖am*	*ǂqx'ann*	*!uin-*	*tsing	*ano-*	*!aan-*

The forms in Table 10 come from geographically adjacent languages of all three families, spreading from the Central Kalahari through northwestern Botswana into northern Namibia. The formal difference between these elements seems to disprove a large amount of diffusion of linguistic material; if contact was involved here, this would rather have been of a conceptual structural nature.[17]

However, the Kalahari Khoe reconstruction *daba~dama conveying "younger" in compound-like terms (in opposition to a presumably inherited form *t(y)a~kya "elder") suggests also occasional borrowing of words, because it is virtually identical with a Proto-Ju form *da'ba~da'ma "child/young of" (Sands 2010: 105). The only segmental difference of the former to the latter is the absence of glottalization on the first vowel - this change can be expected within a borrowing scenario on the part of languages which overall tend to be somewhat simpler in phonological terms.[18]

In the following, we will evaluate the family-internal behavior with respect to the different classification types, also with a view to the geographical patterns observed.

5.2 The Kx'a family

The situation in the Kx'a family is relatively simple. The three attested types of sibling classifications all code relative age and represent a structural continuum whereby Type C in North Ju|'hoan mediates between Type D in all other Ju varieties and Type B in ǂHoan. The least amount of change would be implied by a scenario in which the North Ju|'hoan type is viewed as the proto-system. For ǂHoan, this would imply that it gave up any kind of sex-of-referent distinction - a relatively minor change. In terms of modern contact triggers, Taa may look attractive but is actually equivocal, because the Type-B system need not be old here either (see §5.3). North Ju|'hoan as the most conservative group also seems tempting under the common assumption that it is the most isolated of all San communities who had thus the best opportunity to preserve any old feature. Nevertheless, this explanation has several disadvantages, apart from the fact that this view about them has been strongly debated (cf.

[17] Formally, one could certainly compare ǂHoan *ki* in *ki si(i)* and Naro *kai-*, which is assumed here to derive from Proto-Khoe *kai 'big' but is normally transcribed also as *ki(i)-*. However, the formal similarity could be coincidental and we lack evidence for a plausible borrowing scenario as well as for the precondition that the two languages were in contact in the first place. A very suggestive case of borrowing of linguistic material is, of course, found in Hai‖om as recorded by Widlok (1999), but this would not concern another Khoisan language as source: the Hai‖om forms *ausi-s* "elder sister" and *abudi-b* "elder brother" are most probably recent loans from Afrikaans *ousus* and *(?ou)boetie*, respectively (although Vedder (1923: 162, FN) provides other meanings and etymological derivations for similar words in Damara: *abudi-b* "small father" from *abob* "father" and diminutive *-di*, and *ausi-s* "small mother" from *aos* "wife (honorific)"). Interestingly, parallel to Naro, Hai‖om has innovated the "elder sibling" forms; their "younger sibling" counterpart *!aan-* goes back to the semantically generic Proto-Khoekhoe *!aan- "sibling" and thus must have acquired its specific meaning only secondarily in opposition to the borrowed forms.

[18] There is another candidate for a Ju borrowing in Kalahari Khoe languages. The generic sibling terms *ci-* in Khwe and *si/ce* in |Xokhwe are comparable with the Proto-Ju form *tsing "younger sibling" (see Tables 7 and 3). Ju is the more likely donor in a contact scenario for several reasons: a) *tsing is a reconstructed form with a certain time depth, b) its modern reflexes are phonologically more diverse and complex than the Khoe forms (the latter would result from the simplifying loss of the final nasal of *tsing in the borrowing process), and c) such variants as *tcing* and *cing* in northern Ju dialects are particularly good matches of *ci-* and *si/ce* in the geographically close Kalahari Khoe languages Khwe and |Xokhwe.

Wilmsen 1989). First, North Ju|'hoan is geographically not intermediate between ǂHoan and the rest of Ju, which complicates the historical scenario. Second, its particular intermediate system is an unlikely original stage of a developmental chain (cf. Murdock 1968). Last but not least, there are robust linguistic indications across the different varieties, including North Ju|'hoan itself, that a Type-D system can be reconstructed back to Proto-Ju, implying that North Ju|'hoan changed from Type D to C.

An alternative hypothesis, based on both linguistic and anthropological arguments, is that Proto-Kx'a had a Type-B system and the additional sex-of-referent distinction yielding Type D developed secondarily. These ideas are supported by the phylogenetic analysis for Proto-Kx'a which infers the presence of relative age by means of a lexical term at a probability of 0.97 while the presence of the sex-of-referent distinction is far less certain, with a probability 0.45; Types B and D can be assumed with equal probabilities, while E and F are not supported. While the linguistic evidence for this proposal has been discussed already in §3.2, we add here another idea from an anthropological perspective, viz. that the recognition of relative age seems to be a basic feature of social organization among foragers, at least in the Kalahari Basin: social hierarchy is based on seniority in an otherwise egalitarian society, socialization is an important aspect of the elder/younger sibling relationship, and core groups of siblings (brothers and sisters) are central for band formation and access to resources (cf., e.g., L. Marshall 1976; Lee 1984). This seems to be corroborated by the observation that sibling classifications with a relative-age distinction are found in the area only in communities with a foraging economy. This question will also be relevant in the two other families, and more discussion will follow. So the development of a sex-of-referent distinction in Ju could be a later independent change from a simpler system involving a socially motivated feature to a more complex one. Alternatively or in addition, it could have been influenced by contact with neighboring Khoe languages which possess this feature (see §5.4).

5.3 The Tuu family

As shown in §2 and §3.3, the Tuu family hosts two structurally quite distinct types with entirely different lexical items: Type B in Taa vs. Type E in Lower Nossob and !Ui. This situation is particularly puzzling because on account of other linguistic data (see Güldemann, this volume) Lower Nossob is genealogically closer to Taa and not !Ui. Two principal solutions for this problem can be proposed from a linguistic perspective. One possible scenario is that Lower Nossob used to be like Taa, focusing on relative age, but later acquired the !Ui system, focusing on the sex of referent, thereby borrowing also several linguistic elements (this does not explain *-si in forms for "brother"). Note that this scenario still fails to provide a desirable Proto-Tuu reconstruction. It would, however, be compatible with the intimate contact attested at least between the !Ui language N‖ng and the Lower Nossob language |'Auni.

This scenario also hinges on the possible relationship between a forager economy and a relative-age distinction mentioned above. That is, one might view a trajectory from a Type-B to a Type-E system (losing relative age and gaining sex of referent) as a general trend in the Tuu family induced by contact with encroaching prestige groups speaking Khoekhoe, Bantu[19] and/or Afrikaans for some of which the sex-of-referent distinction is more central. Note that

[19] Possible contact languages from Bantu with terminologies referring to the sex of referent are found in the Nguni group, according to Murdock (1968: 21): Ndebele, Pondo, Swazi, Zulu. Since he (1968: 10) considers the feature to be exceptional in Bantu languages of the wider area, he assumes the borrowing direction to be from Khoisan languages into Nguni. Although this hypothesis is contrary to the general assumption that kinship terminology tends to be imposed by colonizing cultures rather than enter them through substrate interference (Dziebel 2007: 152f), it can certainly not be dismissed in view of a number of other features in Nguni that are due to contact with both Khoekhoe and/or !Ui groups.

the Lower Nossob and !Ui languages are the only attested exceptions to the correlation between foraging and relative-age in the Kalahari Basin.

Modern anthropological data appear to support this view. In the West !Xoon variety of Taa in Namibia a deviant use of sibling terms was observed among a number of young adults who spent several years in boarding schools with multiethnic peer groups. There, a sex-of-referent distinction is the main conceptual dimension in the kin terms of the media of instruction, Afrikaans or English, and is embodied in the everyday experience of hostel life. This environment seems to be responsible for the fact that young Taa speakers, although fluent in the language, use the Taa words for "elder sibling" and "younger sibling" for "brother" and "sister", respectively (Boden 2012).

Analogous shifts may also have occurred historically where small Tuu-speaking communities further south became surrounded by larger groups, eventually shifting to the relevant prestige language. Sociolinguistic information to this effect is amply attested for both Lower Nossob and !Ui languages (cf. Traill 1996, 2002).

An alternative to the hypothesis in terms of innovation on the part of Lower Nossob (and !Ui) is that the Type-E system common to these languages is closer to that of Proto-Tuu, implying that Taa innovated Type B. This scenario has the following advantages: a) it directly provides a Proto-Tuu reconstruction of either Type E or, yet earlier, Type A, together with concrete linguistic forms; b) it requires a change on the part of just one language; and c) given the possibility that Proto-Tuu had Type A, the structural change from this to Type B in Taa and to Type E in !Ui and Lower Nossob would have been quite simple. Moreover, potential language contact would provide a plausible motivation in both cases: for !Ui and Lower Nossob this has been discussed already; for Taa it can be said that it is part of the northern relative-age area and is a direct neighbor of ǂHoan, the only other language with Type-B in the entire region.

Since the new Taa terms are obviously not borrowed, the question arises as to how they were recruited otherwise. For *ǂqx'ann "younger sibling", linguistic clues make it possible that this term continues an inherited word found in Taa's relatives of the !Ui branch. Compare the following forms for "small, little, young" given by Bleek (1956: 643, 652) and Honken (ms.).

(2)	|Xam	*ǂer̃r̃i*	(= /ǂ'aʳni/)	(singular)
		ǂen(ni)	(= /ǂ'ann/~/ǂ'ani/)	(plural)
	N‖ng	*ǂĩ*	(= /ǂ'ain/)	
	‖Xegwi	*kl'ini*	(= /tɬ'ani/)	(also used in "younger sister/brother")

All these words may be related by natural and attested sound changes to a proto-form *ǂ'a$^{(r)}$ni. Bleek (1956: 668) also records another root for "small" in a |Xam compound, *ǂk''erri* (= /ǂkx'ari/), which is most probably related to the previous set of words and suggests a yet earlier Proto-!Ui form *ǂkx'a$^{(r)}$ni. Since this is suspiciously close to Taa *ǂqx'ann, one can hypothesize that Taa re-lexicalized an old stem "small/little/young" for conveying "younger sibling" - partly similar to the case of Naro which is assumed to have recruited the Proto-Khoe adjective *kai "big" for "elder sibling" (cf. §3.4).

The probability of the two contrasting hypotheses is not disambiguated by the PCM analysis. No robust ancestral states could be inferred; sex of referent expressed via a suffix was the dimension with the most support with a 0.66 probability for !Ui and a yet lower value of 0.5 for Proto-Tuu; the relative-age dimension was not supported for Proto-Tuu. Although the PCM analysis did not incorporate the option of a single sibling term, the results for Proto-Tuu are consistent with this notion and modification for sex of referent. Type B (0.33) and especially Type E (0.64) are inferred as the ancestral states, while Types D and F are not supported. With the present knowledge on Tuu history we have to leave it at entertaining these opposite hypotheses and hope that future research will allow a more conclusive solution.

5.4 The Khoe family

The situation in Khoe is the most complex one, in line with previous findings about its diverse internal profile both in terms of linguistic and non-linguistic features (cf. Güldemann 2008). Here, none of the three relevant social dimensions is common to all major groups, nor are there any robust lexical reconstructions.

A relatively simple historical scenario concerns the sex-of-referent dimension in that Proto-Khoe can be said to have had the possibility to mark it by means of the inherited grammatical gender system. This purely linguistic feature of all nouns became more and more obligatory in parts of the family with an effect on the kinship systems, in line with the result of the PCM analysis in §4. The main question for the Proto-Khoe sibling system is whether it expressed relative age, relative sex, both, or neither.

The distribution of an entrenched relative-age distinction in Khoe is as follows: it is present in the north of the Kalahari Basin in Kalahari Khoe as well as most Khoekhoe varieties in Namibia spoken by foragers, viz. Hai‖om, ǂAakhoe, and some Damara groups; it is absent in pastoral Khoekhoe varieties originating in the south.[20] The structural divide within Khoekhoe varieties seems to be particularly suggestive for the validity of the possible correlation between forager subsistence and relative-age distinction in sibling terms. The rise or loss of this feature could even be related to the transition between a forager and a herder society, as claimed for other features of Khoe kinship systems (Barnard 1992). However, one should also consider a somewhat different view; the picture could be an attenuated reflex of the general areal pattern outlined above: absence of a relative-age distinction in the south vs. its presence in the north, whereby the pastoral Khoekhoe retain their "southern" system without relative age but the northern "peripheral" Khoekhoe varieties spoken by foragers align themselves with the areal trend predominating here.

In Kalahari Khoe the situation is different again because the modifiers *t(y)a~kya "elder" and *daba~dama "younger" are dedicated to the use in sibling terms and can be reconstructed, despite the geographical cline according to which they encode the feature only for same sex siblings in the east but the entire system in the west. Even here the relevant social dimension was not as deeply entrenched in earlier stages as it is today in some languages. If the borrowing hypothesis for *daba~dama is corroborated by future research, one could even assume that the introduction of the feature was also due to local areal contact with a Ju-like Kx'a language.

Sibling terms referring to "elder" are conveyed in Khoe by two roots: *t(y)a~kya in most of the Kalahari branch or reflexes of Proto-Khoe *kai 'big' in Naro and facultatively in Khoekhoe.[21] In Naro the adjectival modifier has become the lexical base, indicating a deeper entrenchment of the feature in the terminology system.

---

[20] This finding is related to our decision to disregard adjectives conveying "elder" or "young(er)" and diminutives as elements that are dedicated to the terminology system, although it is not always possible to determine their actual frequency and context of use. Such modifying devices are found in both Nama and Damara dialects. For example, Hoernlé (1985: 53) documents a Nama variety with *sisi-ro-s* "younger sister" and *buti-ro-b* "younger brother" in which "younger" is expressed by the diminutive suffix *-ro*. The lexical stems are transparently loan words from Afrikaans *sussie* "sister" and *boetie* "brother", similar to the Hai‖om case briefly discussed above. Damara varieties with a similar relative-age distinction are recorded by Vedder (1923: 162), Barnard (1992: 208), Haacke & Eiseb (2002), and Schnegg & Pauli (2010: 312).

[21] Whether the two forms are themselves etymologically related is still unclear. The situation must be more complex than assumed by Güldemann and Elderkin (2010: 25) who simply associate Proto-Khoe *kai 'big' with Kwadi *kye(na)* "big/old".

In spite of all the historical uncertainties, it can be generalized that all devices found in the Khoe family to convey relative age are ultimately secondary so that the proto-system would, if anything, have marked the feature facultatively, similar to the situation in pastoral Khoekhoe varieties. Its later establishment as an obligatory feature in parts of the family is again supported by the PCM analysis of §4.

The historical evaluation of the relative-sex distinction turns out to be the most equivocal. It is only found in certain Kalahari Khoe languages in the northeast, Shua, Ganadi and |Xokhwe, which might suggest that it is a local innovation. Since these groups were in close contact with Bantu groups who also possess this feature, structural borrowing on the part of these Khoe languages would thus be one possible explanation, all the more so since these groups have been described as "acculturated" to Bantu (Barnard 1992: 121).

It is noteworthy that their terms for "opposite sex sibling" are probably cognate with the Khwe form *kŭĭ-kx'ei* (cf. Kilian-Hatz (2003: 64, 251), transliterated here as *kuin-kx'ei*), meaning "taboo for joking and marriage" and used for "opposite sex parallel cousin" as well as, if rarely, for "opposite sex cross cousin of first degree" (cf. Boden 2005, pace Kilian-Hatz who translates it just as "mother's brother's child"). In an earlier stage, *kuin-kx'ei* might have had the general meaning "opposite sex or avoidance cousin". Hence, it is conceivable that the cognate of *kuin-kx'ei* in eastern Kalahari Khoe had the same meaning but was recruited for "opposite sex sibling" when the languages took over the Type-F classification from Bantu. In Khwe the shift would have taken an opposite direction: since its term for "cross cousin", *cìroó*, is a loan from Mbukushu *thiro* (cf. Larson 1977), *kuin-kx'ei* would have come to be used, if at all, for second degree cousins of opposite sex.

An alternative hypothesis for the dynamics of the relative-sex dimension in Khoe would be to assume that it is original but, in line with the PCM analysis, was lost in the western and southern areas of the family. Geographically, the loss of the feature could be located already at the boundary between East and West Kalahari Khoe.

This boundary may in fact be of more general importance for the dynamics of sibling terminologies in Khoe, epitomized by |Xokhwe, which is classified linguistically with the West branch but whose sibling terminology is of the same type as in the East branch. Under the possible hypothesis that the system in the (north)east is conservative in the family, |Xokhwe would represent a kind of family-internal pivot where crucial systemic transitions take place: namely from a system with a robust lexicalized relative-sex distinction, a weaker relative-age distinction, and a somewhat marginal sex-of-referent distinction towards a western system with an absent relative-sex distinction, a consistent relative-age distinction, and a more grammaticalized sex-of-referent distinction. In view of the linguistically transparent and thus presumably later emergence of relative age by means of *t(y)a~kya and *daba~dama, it is even conceivable that the earliest stage in Khoe only had relative sex as an obligatory feature, which, however, is not attested anymore synchronically.

The idea of northeastern Khoe conservatism also relates to possible dynamics further south in the Khoekhoe branch. This is directly associated with the other large sibling-terminology area in the southern part of the Kalahari Basin, which is in virtually complementary geographical distribution with the northern relative-age area. A common denominator of the southern area is the existence of simplex lexical items with a securely reconstructable meaning "sibling", as given in Table 11; this feature even justifies the hypothesis that at least some Type-E systems go back to earlier and simpler Type-A systems.[22]

[22] There are also generic sibling forms in West Kalahari Khoe languages other than Naro (see Tables 7 and 8). However, as discussed in §3.4, these are unclear in terms of their semantic reconstruction and/or origin and thus should not be considered here.

Table 11: Reconstructed forms for generic “sibling” across families

Meaning	Lower Nossob (Tuu)	!Ui (Tuu)	Khoekhoe (Khoe)
Sibling	*‖aa	*‖aa(n)	*!aan

Disregarding the later North Khoekhoe expansion which brought Type E to the northwest, the generic-sibling region coincides neatly with the Cape linguistic area proposed by Güldemann (2006) which is argued to have come about to a large extent by substrate interference from Tuu languages in colonizing Khoekhoe. In line with such a historical scenario, it is worth considering that the relevant sibling-terminology feature in Proto-Khoekhoe is a contact-induced innovation,[23] opposed to Kalahari Khoe where dedicated lexical terms focus on relative sex and/or relative age.

Overall, the idea of reconstructing the Proto-Khoe system for sibling terms as Type F, which focuses on relative sex, and assuming that Khoe languages further west and south innovated under contact pressure is compatible with a more general historical scenario for the family proposed by Güldemann (2008). In particular, the individual innovations can be motivated by the relevant areal environment in that the assumed indigenous contact languages provided the new features: relative-age in the north from Type B or D in Kx’a vs. a generic sibling term in the south from Type E or A in southern Tuu. This scenario would also allow one to project back some modern linguistic elements to Proto-Khoe, which is not possible under the alternative scenario, particularly *xo “same sex sibling” and *!uin-(k(’)e) “opposite sex sibling”.

However, we are still far from being able to provide a more conclusive history of the sibling terminology in Khoe. The PCM analysis also cannot help here. While it supports a relative-age distinction for Proto-Kalahari Khoe with a probability of 0.75 and a Type E system for Proto-Khoekhoe with a probability of 0.64, all other inferred types and dimensions are equivocal at best. The major alternatives for Khoe as a whole revolve primarily around the status of the relative-sex distinction. An answer to the question which hypothesis is correct has to be deferred until more is known about the general history of the area. Here and elsewhere, it is crucial in the future to include fuller data about the systems in Bantu languages in order to determine the more likely direction of influence in case of culture and language contact.

6. General conclusions

Based on anthropological comparison, historical linguistics, and statistically relevant PCM, we developed specific hypotheses about the historical dynamics in the sibling terminologies of Khoisan languages. Several general conclusions can be drawn for the historical questions in the Kalahari Basin and this type of historical research in general.

First, while Barnard (particularly 1988, 1992) identified common features of kinship systems across Khoisan, we could not find any shared deep features in the sibling terminologies analyzed here, let alone vocabulary shared across the major groups and derived from a common ancestor system. Instead the most likely scenario for this kinship domain is one where the three language families are independent lineages in which sibling classification evolved along different trajectories. This major finding conforms to the current specialist view that Khoisan is an areal entity at best.

The combination of anthropological, linguistic, and statistical analyses was also applied to the reconstruction of family-internal processes. This, we argue, achieved superior results than

[23] This begs the question whether *!aan in Khoekhoe is itself a borrowing, related to the quite similar *‖aa(n) in Tuu. Phonetically both could come from one and the same source, notably *!!aan. This would imply that Khoekhoe borrowed a masculine form with nasalization but a vowel sequence /aa/ (like in |Xam) from a !Ui variety where the sound change from /!!/ to /‖/ had not (yet) taken place. A major problem of this idea is that /!!/ has so far not at all been attested or assumed for any language in Tuu (and Khoe for that matter).

would have been possible within any one discipline alone, if only for the fact that such an approach helps to evaluate a situation where different disciplines produce competing hypotheses. Our major hypotheses are summarized in (3)-(5) (an arrow of the form -[*Family*]-> indicates that the relevant change may have been influenced by contact with the family/ies within the square brackets).

(3) Kx'a
B (Proto-Kx'a, ǂHoan) -[*Khoe*]-> D (Ju) --> C (Ju|'hoan)

(4) Tuu
a. B (Proto-Tuu, Taa) -[*Khoekhoe, Bantu, Germanic*]-> E (!Ui, Lower Nossob)
b. A (Proto-Tuu) -[*Khoekhoe, Bantu, Germanic*]-> E (!Ui, Lower Nossob)
-[*Kx'a*]-> B (Taa)

(5) Khoe
a. ? (Proto-Khoe) --> D (West Kalahari) -[*Bantu*]-> F (East Kalahari)
-[*Tuu*]-> E (Khoekhoe) -[*Kx'a*]-> D (Hai‖om, ǂAakhoe)
b. F_1 (Proto-Khoe) -[*Kx'a*]-> F_2 (East Kalahari) -[*Kx'a*]-> D (West Kalahari)[24]
-[*Tuu*]-> E (Khoekhoe) -[*Kx'a*]-> D (Hai‖om, ǂAakhoe)

In general, the picture for the Kx'a family in (3) is relatively the least controversial, while the data available for Tuu and Khoe allow for at least two, partly opposite historical reconstructions, as shown in (4) and (5), respectively. The alternative scenarios here are mainly due to differences in the assumed source and direction concerning the probable cultural borrowing that seems to have affected kinship classifications in the area, particularly the relative role attributed to Non-Khoisan languages, particularly from Germanic and Bantu, on the one hand (more prominent in the a.-scenarios) and Khoisan languages on the other hand (more prominent in the b.-scenarios). The insufficient coverage of Bantu within the first group is a major drawback in our present analysis and removing this gap promises to clarify at least some of the open problems.

Whatever the final verdict, the general importance of culture and language contact in shaping the modern profile of sibling terminologies in the Khoisan domain seems to be beyond doubt. Historical scenarios of this kind help explain family-internal mismatches between genealogical affiliation and sibling terminology system (cf. the most obvious cases of |Xokhwe within Kalahari Khoe and Lower Nossob within Tuu), and can also motivate the existence of large-scale areal patterns that cross family boundaries (notably the relative-age area in the north and the generic-sibling area in the south, cf. Map 1 above).

The results of our study also throw new light on the general discussion concerning the historical dynamics of kinship systems. First, they show that sibling classifications can shift independently of lexical changes and do not necessarily involve borrowing of linguistic substance (pace Blust 1994). Second and more importantly, they cast doubt on the conservatism of kinship terminology in general and sibling classification in particular, as assumed, for example, by Trautmann (2008). Moreover, kinship systems are not only subject to universal processes but also to family- and/or area-specific trajectories of change.

References

Barnard, Alan. 1976. Nharo Bushman Kinship and the Transformation of Khoi Kin Categories. Ph.D. dissertation. University College London.

[24] The difference between F_1 and F_2 lies in the absence or presence of coding for relative age.

Barnard, Alan. 1985. A Nharo Wordlist with Notes on Grammar. Durban: University of Natal.
Barnard, Alan. 1988. Kinship, Language and Production: A Conjectural History of Khoisan Social Structure. Africa 18.216-236.
Barnard, Alan. 1992. Hunters and Herders of Southern Africa: A Comparative Ethnography of the Khoisan People. (=Cambridge Studies in Social and Cultural Anthropology, 5.) Cambridge: Cambridge University Press.
Bleek, Dorothea F. 1924. Bushman Terms of Relationship. Bantu Studies 2.57-70.
Bleek, Dorothea F. 1929. Comparative Vocabularies of Bushman Languages. Cambridge: Cambridge University Press.
Bleek, Dorothea F. 1937. Grammatical notes and texts in the |Auni language, |Auni vocabulary. In Rheinallt Jones and Doke (eds.), 195-220.
Bleek, Dorothea F. 1956. A Bushman Dictionary. New Haven: American Oriental Society.
Blust, Robert. 1994. Austronesian Sibling Terms and Culture History. Austronesian Terminologies: Continuity and Change ed. by Pawley, Andrew & Malcolm Ross, 31-72. Canberra: Australian National University.
Boden, Gertrud. 2005. Prozesse sozialen Wandels vor dem Hintergrund staatlicher Eingriffe. Eine Fallstudie zu den Khwe in West Caprivi/Namibia. Ph.D. dissertation. University of Cologne.
Boden, Gertrud. 2012. Flexible, Universal and Non-genealogical: Changing Kinship Classifications of San living in Multilingual Social Environments. Paper presented at the 50th anniversary conference of the Centre of African Studies, Edinburgh, June 2012.
Borgerhoff Mulder, Monique, Charles L. Nunn, & Mary. C. Towner. 2006. Cultural macroevolution and the transmission of traits. Evolutionary Anthropology 15:2. 52-64.
Brenzinger, Matthias and Christa König (eds.). 2010. Khoisan languages and linguistics: proceedings of the 1st International Symposium January 4-8, 2003, Riezlern/Kleinwalsertal. Quellen zur Khoisan-Forschung 24. Köln: Rüdiger Köppe.
Budack, Kuno F. R. 1986. Die Klassifikation der Khwe-khwen (Naman) in Südwestafrika. In Voßen, Rainer and Klaus Keuthmann (eds.), Contemporary studies on Khoisan, 2 vols. Quellen zur Khoisan-Forschung 5. Hamburg: Helmut Buske, vol. 1: 107-143.
Dickens, Patrick. 1994. English - Ju|'hoan, Ju|'hoan - English Dictionary. Cologne: Rüdiger Köppe.
Dziebel, German V. 2007. The Genius of Kinship: The Phenomenon of Human Kinship and the Global Diversity of Kinship Terminologies. Youngstown, NY: Cambria Press.
Ehret, Christopher. 2008. Reconstructing Ancient Kinship in Africa. Early Human Kinship. From Sex to Social Reproduction ed. by Allen, Nicholas J., Hillary Callan, Robin Dunbar & Wendy James, 200-231 & Appendix, 259-269. Malden, MA & Oxford: Blackwell Publishing.
Elmendorf, William W. 1961. System Change in Salish Kinship Terminologies. Southwestern Journal of Anthropology 17:4.365-382.
Engelbrecht, Jan A. 1936. The Korana: An Account of their Customs and their History, with Texts. Cape Town: Maskew Miller.
Epling, P. J., Jerome Kirk & John P. Boyd. 1973. Genetic Relations of Polynesian Sibling Terminologies. American Anthropologist 75:5.1596-1625.
Firth, Raymond. 1970. Sibling terms in Polynesia. Journal of the Polynesian Society 79:3.272-287.
Gray, Russell D. & Quentin D. Atkinson. 2003. Language-tree divergence times support the Anatolian theory of Indo-European origin. Nature 426 (6965):435-439.
Gray, Russell D., Alexei J. Drummond & Simon J. Greenhill. 2009. Language Phylogenies Reveal Expansion Pulses and Pauses in Pacific Settlement. Science 323:5913. 479-483.
Gray, Russell D., Simon J. Greenhill & Robert M. Ross. 2007. The Pleasures and Perils of Darwinizing Culture (with Phylogenies). Biological Theory 2:4. 360-375.
Gruber, Jeffrey S. 1973. ǂHoa Kinship Terms. Linguistic Inquiry IV:4.427-449.
Güldemann, Tom. 2004. Reconstruction through 'de-construction': the marking of person, gender, and number in the Khoe family and Kwadi. Diachronica 21,2: 251-306.
Güldemann, Tom. 2005. Studies in Tuu (Southern Khoisan). (= University of Leipzig Papers on Africa, Languages and Literatures 23.) Leipzig: University of Leipzig, Institute of African Studies.
Güldemann, Tom. 2006. Structural Isoglosses between Khoekhoe and Tuu: The Cape as a Linguistic Area. Linguistic Areas: Convergence in Historical and Typological Perspective ed. by Matras, Yaron, April McMahon and Nigel Vincent, 99-134. Hampshire: Palgrave Macmillan.

Güldemann, Tom. 2008. A Linguist's View: Khoe-Kwadi Speakers as the Earliest Food-producers of Southern Africa. Southern African Humanities 20.93-132.
Güldemann, Tom. forthcoming. Person-gender-number marking from Proto-Khoe-Kwadi to its descendents: a rejoinder with particular reference to language contact. In König, Christa and Rainer Voßen (eds.), Festschrift for Bernd Heine. Routledge African Linguistics Series. London: Routledge.
Güldemann, Tom & Edward D. Elderkin. 2010. On external genealogical relationships of the Khoe family. In Brenzinger and König (eds.), 15-52.
Haacke, Wilfried H. G. & Eliphas Eiseb. 2002. A Khoekhoegowab Dictionary with an English-Khoekhoegowab Index. Windhoek; Gamsberg Macmillan.
Hage, Per. 1999. Marking Universals and the Structure and Evolution of Kinship Terminologies. Evidence from Salish. Journal of the Royal Anthropological Institute (NS) 5:3.423-441.
Harmon, Luke J., Jason T. Weir, Chad D. Brock, Richard E. Glor & Wendell Challenger. 2008. GEIGER: investigating evolutionary radiations. Bioinformatics 24:1. 129-131.
Hedican, Edward J. 1986. The Evolution of Apache-Athapaskan Sibling Terminolgy. A Reconsideration. Western Canadian Anthropologist 3.1-18.
Heine, Bernd & Henry Honken. 2010. The Kx'a Family. A New Khoisan Genealogy. Journal of Asian and African Studies 79:3.5-36.
Heinz, Hans-Joachim. n.d. The /xokwe Bugakwe. Unpublished Manuscript.
Heinz, Hans-Joachim. 1994 [1966]. Social Organization of the !Ko Bushmen ed. by Klaus Keuthmann (= Research in Khoisan Studies 10). Cologne: Rüdiger Köppe.
Hoernlé, Winifred A. 1985 [1925]. The Social Organization of the Nama Hottentots of Southwest Africa and Other Essays ed. by Peter Carstens. Johannesburg: Witwatersrand University Press.
Holden, Clare. J. 2002. Bantu language trees reflect the spread of farming across sub- Saharan Africa: a maximum-parsimony analysis. Proceedings of the Royal Society B: Biological Sciences 269:1493. 793-799.
Honken, Henry. ms. Short grammar and dictionary of ǁXegwi. unpublished
Jordan, Fiona M. 2011. A Phylogenetic Analysis of the Evolution of Austronesian Sibling Terminologies. Human Biology 83:2.297-321.
Jordan, Fiona M. forthcoming. Comparative phylogenetic methods and the study of pattern and process in kinship. In Patrick McConvell, Ian Keen & Rachel Hendery (eds) Kinship Systems: Change and Reconstruction. Salt Lake City: University of Utah Press.
Kilian-Hatz, Christa. 2003. Khwe Dictionary with a Supplement on Khwe Place-Names of West Caprivi by Matthias Brenzinger. (=Namibian African Studies 7). Cologne: Rüdiger Köppe.
Kitchen, Andrew, Christopher Ehret, Shiferaw Assefa, & Connie J. Mulligan. 2009. Bayesian phylogenetic analysis of Semitic languages identifies an Early Bronze Age origin of Semitic in the Near East. Proceedings of the Royal Society B: Biological Sciences 276:1668. 2703-2710.
Klocke-Daffa, Sabine. 2001. 'Wenn Du hast, mußt Du geben.' Soziale Sicherung im Ritus und im Alltag bei den Nama von Berseba/Namibia. Münster: LIT.
Kroeber, Alfred L. 1909. Classificatory Systems of Relationship. Journal of the Royal Anthropological Institute 39.77-84.
Kuper, Adam. 1987. South Africa and the Anthropologist. London & New York: Routledge & Kegan Paul.
Larson, Thomas J. 1977. Kinship Terminology of the Hambukushu of Ngamiland. Botswana Notes and Records 9.85-89.
Lee, Richard B. 1984. The Dobe !Kung. New York: Holt, Rinehart & Winston.
Levinson, Stephen C. & Russell D. Gray. 2012. Tools from evolutionary biology shed new light on the diversification of languages. Trends in Cognitive Sciences 16:3. 167-173.
Mace, Ruth & Fiona M. Jordan. 2011. Macro-evolutionary studies of cultural diversity: a review of empirical studies of cultural transmission and cultural adaptation. Philosophical Transactions of the Royal Society B: Biological Sciences 366:1563. 402-411.
Maddison, Wayne. P. & David.R. Maddison. 2011. Mesquite: a modular system for evolutionary analysis. Version 2.75 http://mesquiteproject.org
Maingard, Louis F. 1937. The ǂKhomani dialect of Bushman, its morphology and other characteristics. In Rheinallt Jones and Doke (eds.), 237-275.
Marshall, Lorna. 1957. The Kin Terminology System of the !Kung Bushmen. Africa 27.1-25.

Marshall, Lorna. 1976. The !Kung of Nyae Nyae. Cambridge, Ma., Harvard University Press.
Marshall, Mac. 1984. Structural Patterns of Sibling Classification in Island Oceania: Implications for culture history. Current Anthropolgy 25:5.597-637.
Morgan, Lewis H. 1871. Systems of Consanguinity and Affinity of the Human Family. (=Smithsonian Contributions to Knowledge XVII=218). Washington: Smithsonian Institute.
Murdock, George P. 1949. Social Structure. New York: Macmillan
Murdock, George P. 1968. Patterns of sibling terminology. Ethnology 7.1-24.
Nunn, C. L. 2011. The Comparative Method in Evolutionary Anthropology and Biology. Chicago: University of Chicago Press.
Ono, Hitomi. 1996. An Ethnosemantic Analysis of |Gui Relationship Terminology. African Study Monographs Suppl. 22, 125-144. Kyoto: The Center of African Area Studies.
Ono, Hitomi. 1997. How do the |Gui Categorize People?. Gengo Kenkyu Journal of the Linguistic Society of Japan 111.42-57.
Pagel, Mark, Andrew Meade & Daniel Barker. 2004. Bayesian Estimation of Ancestral Character States on Phylogenies. Systematic Biology 53:5. 673-684.
Paradis, Emmanuel, J. Claude & K. Strimmer. 2004. APE: Analyses of Phylogenetics and Evolution in R language. Bioinformatics 20 (2):289-290.
Potgieter, Evert F. 1955. The Disappearing Bushmen of Lake Chrissie: A Preliminary Survey. Pretoria: J.L. van Schaik.
R Core Team. 2012. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, http://www.R-project.org.
Rheinallt Jones, John D. and Clement M. Doke (eds.). 1937. Bushmen of the Southern Kalahari. Johannesburg: Witwatersrand University Press.
Sands, Bonny E. 2010. Juu subgroups based on phonological patterns. In Brenzinger and König (eds.), 85-114.
Schnegg, Michael & Julia Pauli. 2010. Namibische Wahlverwandtschaften: Zur Dualität von Struktur und Handlungsfreiheit in der Verwandtschaftsethnologie. Verwandtschaft heute. Positionen, Ergebnisse und Perpektiven. ed. by Alber, Erdmute, Bettina Beer, Julia Pauli & Michael Schnegg, 305-328. Berlin: Dietrich Reimer.
Silberbauer, George B. 1981. Hunter and Habitat in the Central Kalahari Desert. Cambridge: Cambridge University Press.
Takada, Akira. 2008. Kinship and Naming among the Ekoka !Xun. Proceedings of the 2nd International Symposium on Khoisan Languages and Linguistics, January 8-12, 2006, Riezlern/Kleinwalsertal ed. by Ermisch, Sonja, 303-322. Cologne: Rüdiger Köppe.
Tanaka, Jiro. 1978. A San Vocabulary of the Central Kalahari. G‖ana and G|wi Dialects. Tokyo: ILCAA University of Foreign Studies.
Tanaka, Jiro. 1980. The San Hunter-Gatherers of the Kalahari : A Study in Ecological Anthropology. Tokyo: University of Tokyo Press.
Traill, Anthony. 1994. A !Xóõ Dictionary. (= Research in Khoisan Studies 9). Cologne: Rüdiger Köppe.
Traill, Anthony. 1996. !Khwa-ka hhouiten hhouiten, 'The rush of the storm': the linguistic death of |Xam. In Skotnes, Pippa (ed.), Miscast: negotiating the presence of the Bushmen. Cape Town: Cape Town University Press, 161-183.
Traill, Anthony. 1999. K'u|ha:si Manuscript by Robert Story. (= Khoisan Forum Working Paper 13). Cologne: University of Cologne SFB 389, Institute of African Studies.
Traill, Anthony. 2002. The Khoesan languages. In Mesthrie, Rajend (ed.), Language in South Africa. Cambridge: Cambridge University Press, 27-49.
Trautmann, Thomas R. 2008. The Whole History of Kinship Terminologies in Three Chapters: Before Morgan, Morgan, and After Morgan. Lewis Henry Morgan and the Invention of Kinship. ed. by Trautmann, Thomas R., 308-333. Lincoln & London: University of Nebraska Press.
Vedder, Heinrich. 1923. Die Bergdama. 2 vols. Hamburg: L. Friedrichsen & Co.
Vedder, Heinrich. 1997 [1934]. Das alte Südwestafrika: Südwestafrikas Geschichte bis zum Tode Mahareros 1890. Windhoek: Namibia Wissenschaftliche Gesellschaft.
Visser, Hessel. 2001. Naro Dictionary: Naro-English/ English-Naro. Gantsi: Naro Language Project.
Voßen, Rainer. 1997. Die Khoe-Sprachen. Ein Beitrag zur Sprachgeschichte Afrikas. Cologne: Rüdiger Köppe.

Westphal, Ernst O. J. n.d. Ganade papers. African Studies Archive, University of Cape Town.
Widlok, Thomas. 1999. Living on Mangetti. Bushman Autonomy and Namibian Independence. Oxford: Oxford University Press.
Wilmsen, Edwin N. 1989. Land Filled with Flies: A Political Economy of the Kalahari. Chicago: University of Chicago Press.

Table 1: A survey of sibling classifications across Khoisan

No	Language (variety)	Genealogical classification	Type	Terms for	Subsistence	Selected sources
Khoe-Kwadi						
1	!Ora	Khoe, Khoekhoe, South	E	B, Z	pastoralist	Engelbrecht 1936
2	Nama	Khoe, Khoekhoe, North	E	B, Z	pastoralist	Haacke & Eiseb 2002, Hoernlé 1985, Klocke-Daffa 2001
3	Damara	Khoe, Khoekhoe, North	E	B, Z	foraging	Haacke & Eiseb 2002, Vedder 1923
4	Hai‖om	Khoe, Khoekhoe, North	D	eB, eZ, yB, yZ	foraging	Widlok 1999
5	ǂAakhoe	Khoe, Khoekhoe, North	D	eB, eZ, yB, yZ	foraging	Widlok 1999
6	Shua	Khoe, Kalahari, East	F	essG, yssG, osG	foraging	McGregor & Kure field notes 2010-12
7	Ganadi	Khoe, Kalahari, East	F	essG, yssG, osG	foraging	Westphal field notes 1953?, 1961?
8	ǀXokhwe	Khoe, Kalahari, West	F	essG, yssG, osG	foraging	Heinz n.d., field notes 1970s
9	Khwe	Khoe, Kalahari, West	D	eB, eZ, yB, yZ	foraging	Boden 2005, Kilian-Hatz 2003
10	G‖ana	Khoe, Kalahari, West	D	eB, eZ, yB, yZ	foraging	Tanaka 1978, 1980
11	Gǀui	Khoe, Kalahari, West	D	eB, eZ, yB, yZ	foraging	Ono 1996, 1997; Silberbauer 1981
12	Naro	Khoe, Kalahari, West	D	eB, eZ, yB, yZ	foraging	Barnard 1976, 1985, 1992; Visser 2001
Kx'a						
13	Angola !Xuun	Kx'a, Ju,-North	D	eB, eZ, yB, yZ	foraging	Bleek 1956, Boden field notes 2012
14	Ekoka !Xuun	Kx'a, Ju,-North-central	D	eB, eZ, yB, yZ	foraging	Takada 2008
15	Lloyd's !Xuun	Kx'a, Ju,-North-central	D	eB, eZ, yB, yZ	foraging	Bleek 1924, 1956
16	North Juǀ'hoan	Kx'a, Ju,-Southeast	C	eB, eZ, yG	foraging	Dickens 1994, L. Marshall 1957, Lee 1984
17	South Juǀ'hoan	Kx'a, Ju,-Southeast	D	eB, eZ, yB, yZ	foraging	Bleek 1929, Pratchett & Boden field notes 2011
18	ǂHoan	Kx'a	B	eG, yG	foraging	Gruber 1973
Tuu						
19	Taa	Tuu, Taa-Lower Nossob	B	eG, yG	foraging	Heinz 1994, Traill 1994
20	ǀ'Auni	Tuu, Taa-Lower Nossob	E	B, Z	foraging	Bleek 1937
21	ǀHaasi	Tuu, Taa-Lower Nossob	E	B, Z	foraging	Traill 1999
22	N‖ng	Tuu, !Ui	E	B, Z	foraging	Bleek 1929, 1956; Maingard 1937
23	ǀXam	Tuu, !Ui	E	B, Z	foraging	Bleek 1924, 1956
24	‖Xegwi	Tuu, !Ui	E	B, Z	foraging	Potgieter 1955, Honken ms.